Gastroenterology 2016;150:626–637

CLINICAL—LIVER

Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography

Kento Imajo,[1] Takaomi Kessoku,[1] Yasushi Honda,[1] Wataru Tomeno,[1] Yuji Ogawa,[1] Hironori Mawatari,[1] Koji Fujita,[1] Masato Yoneda,[1] Masataka Taguri,[2] Hideyuki Hyogo,[3] Yoshio Sumida,[4] Masafumi Ono,[5] Yuichiro Eguchi,[6] Tomio Inoue,[7] Takeharu Yamanaka,[2] Koichiro Wada,[8] Satoru Saito,[1] and Atsushi Nakajima[1]

[1]Department of Gastroenterology, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [2]Department of Biostatistics and Epidemiology, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [3]Department of Medicine and Molecular Science, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima, Japan; [4]Department of Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, Kamigyo-ku, Kyoto, Japan; [5]Department of Gastroenterology and Hepatology, Kochi Medical School, Kochi, Japan; [6]Division of Hepatology, Saga Medical School, Liver Center, Saga, Japan; [7]Department of Radiology, Yokohama City University Graduate School of Medicine, Yokohama, Japan; and [8]Department of Pharmacology, Shimane University Faculty of Medicine, Izumo, Shimane, Japan

See Covering the Cover synopsis on page 540; see editorial on page 553.

BACKGROUND & AIMS: Noninvasive methods have been evaluated for the assessment of liver fibrosis and steatosis in patients with nonalcoholic fatty liver disease (NAFLD). We compared the ability of transient elastography (TE) with the M-probe, and magnetic resonance elastography (MRE) to assess liver fibrosis. Findings from magnetic resonance imaging (MRI)–based proton density fat fraction (PDFF) measurements were compared with those from TE-based controlled attenuation parameter (CAP) measurements to assess steatosis. **METHODS:** We performed a cross-sectional study of 142 patients with NAFLD (identified by liver biopsy; mean body mass index, 28.1 kg/m^2) in Japan from July 2013 through April 2015. Our study also included 10 comparable subjects without NAFLD (controls). All study subjects were evaluated by TE (including CAP measurements), MRI using the MRE and PDFF techniques. **RESULTS:** TE identified patients with fibrosis stage $\geq$2 with an area under the receiver operating characteristic (AUROC) curve value of 0.82 (95% confidence interval [CI]: 0.74–0.89), whereas MRE identified these patients with an AUROC curve value of 0.91 (95% CI: 0.86–0.96; $P = .001$). TE-based CAP measurements identified patients with hepatic steatosis grade $\geq$2 with an AUROC curve value of 0.73 (95% CI: 0.64–0.81) and PDFF methods identified them with an AUROC curve value of 0.90 (95% CI: 0.82–0.97; $P < .001$). Measurement of serum keratin 18 fragments or alanine aminotransferase did not add value to TE or MRI for identifying nonalcoholic steatohepatitis. **CONCLUSIONS:** MRE and PDFF methods have higher diagnostic performance in noninvasive detection of liver fibrosis and steatosis in patients with NAFLD than TE and CAP methods. MRI-based noninvasive assessment of liver fibrosis and steatosis is a potential alternative to liver biopsy in clinical practice. UMIN Clinical Trials Registry No. UMIN000012757.

Keywords: Diagnosis; Classification; Alanine Transaminase; Overweight.

Nonalcoholic fatty liver disease (NAFLD) is an important cause of chronic liver injury in many countries.[1,2] A recent study has shown that the risk of developing NAFLD is 4–11 times higher in patients with metabolic syndrome as compared with healthy individuals.[3] NAFLD ranges from benign nonalcoholic fatty liver (NAFL) to nonalcoholic steatohepatitis (NASH). This latter condition includes progressive fibrosis[4] and hepatocellular carcinoma.[5,6] Liver biopsy is recommended as the gold standard method for the diagnosis and grading of steatosis, hepatic inflammation, and hepatocellular ballooning, and the staging of liver fibrosis in patients with NASH.[7] However, because of increased cost, possible risk, and health care resource use, an invasive liver biopsy is a poorly suited diagnostic test for such a prevalent condition.[8] In addition, the histologic lesions of NASH are unevenly distributed throughout the liver parenchyma, therefore, liver biopsy sampling error can result in substantial stratification and staging inaccuracies.[9]

Assessment of the severity of liver fibrosis and steatosis is important in the evaluation of the stage of NAFLD. Transient elastography (TE; Fibroscan, EchoSens, Paris, France) is a useful technique that allows rapid and noninvasive measurement of mean tissue stiffness.[10] We have reported that liver stiffness measurement (LSM) obtained using TE was useful for estimation of severity of liver fibrosis in NAFLD.[11,12] Meta-analysis of TE has also shown that LSM accurately reflected liver fibrosis.[13] In addition,

Abbreviations used in this paper: **AUROC, area under the receiver operating characteristic; BMI, body mass index; CAP, controlled attenuation parameter; CI, confidence interval; LSM, liver stiffness measurement; MRE, magnetic resonance elastography; MRI, magnetic resonance imaging; NAFL, nonalcoholic fatty liver; NAFLD, nonalcoholic fatty liver disease; NAS, Nonalcoholic Fatty Liver Disease Activity Score; NASH, nonalcoholic steatohepatitis; PDFF, proton density fat fraction; ROI, region of interest; TE, transient elastography.**

Most current article

0016-5085
http://dx.doi.org/10.1053/j.gastro.2015.11.048

recent interest has shifted toward the controlled attenuation parameter (CAP), which is based on the properties of ultrasonic signals acquired using TE.[14] Recent clinical studies using the TE system have demonstrated an increase in the CAP with severe fat accumulation and liver stiffness with advanced fibrosis in NAFLD patients.[15–17] Body mass index (BMI) >28 kg/m^2 has been identified as an independent risk factor for failure to measure liver elastography and fat accumulation.[18] Therefore, TE has the limitation that sometimes LSM and CAP cannot be measured in morbidly obese patients with NAFLD.

Magnetic resonance elastography (MRE) is a magnetic resonance imaging (MRI)-based method for quantitatively imaging tissue stiffness, and is available from several manufacturers of MRI scanners as an option that includes special hardware and software. Quantitative stiffness images (elastograms) of the liver can be obtained rapidly during breath-hold acquisitions, and can therefore be readily included in conventional liver MRI protocols.[19] Multiple studies have also shown that MRE-based LSM provides an accurate biomarker for detecting the presence of fibrosis in patients with chronic liver dysfunction.[20–24] Indeed, MRE has been reported to be a useful method for the diagnosis of liver fibrosis in patients with NAFLD, even in the early stages.[25–27] Proton density fat fraction (PDFF) measurement is an MRI-based method for quantitatively assessing hepatic steatosis and is available from several manufacturers of MRI scanners as an option. MRI-determined PDFF correlates with histologically determined steatosis grade in patients with NAFLD.[28,29] Although MRI can be performed even in morbidly obese patients with NAFLD, where it is sometimes difficult to perform TE, there are no reports that have directly compared the diagnostic accuracy of MRI and TE for assessing both fibrosis and steatosis in patients with NAFLD.

Table 1. Clinical, Serologic, and Histologic Characteristics of Control Subjects and Patients With NAFLD

Characteristic	Control	NAFLD	*P* value
n	10	142	
Age, y, mean ± SD	52.1 ± 15.1	57.5 ± 14.6	.362
Sex, male/female	6/4	81/61	.321
BMI, *kg/m^2*, mean ± SD	21.9 ± 0.69	28.1 ± 4.63	<.001
Platelets, */10^4 μL*, mean ± SD	22.8 ± 4.31	20.9 ± 7.69	.442
AST, *IU/L*, mean ± SD	23.4 ± 8.12	44.5 ± 26.3	<.001
ALT, *IU/L*, mean ± SD	24.3 ± 7.31	56.2 ± 42.6	<.001
γ-GTP, *IU/L*, mean ± SD	39.0 ± 6.73	80.0 ± 87.7	.297
C-reactive protein, *mg/L*, mean ± SD	0.06 ± 0.03	0.17 ± 0.09	.003
Creatinine, *mg/dL*, mean ± SD	0.63 ± 0.32	0.77 ± 0.42	.672
Fasting blood glucose, *mg/dL*, mean ± SD	92.1 ± 16.3	110.3 ± 28.2	.001
Fasting insulin, *μU/mL*, mean ± SD	7.47 ± 2.98	19.2 ± 20.8	.001
HbA1c, mean ± SD	5.62 ± 0.61	6.43 ± 1.10	.002
Diabetes mellitus, %	0	71 (50.0)	
Hypertension, %	0	45 (31.7)	
Dyslipidemia, %	0	94 (66.2)	
Length of specimens, *mm*, mean ± SD		21.3 ± 1.94	
Number of portal areas, mean ± SD		14.3 ± 4.42	
Steatosis grade, n			
5%–33%		59	
33%–66%		59	
>66%		24	
Lobular inflammation (n)			
None		6	
<2 foci per 200× field		78	
2–4 foci per 200× field		52	
>4 foci per 200× field		6	
Liver cell ballooning (n)			
None		32	
Few balloon cells		96	
Many balloon cells		14	
NAFL/NASH, n		34/108	
NAS, n			
1/2/3/4/5/6/7		6/15/32/51/30/5/3	
Fibrosis stage, n			
None		14	
Perisinusoidal or periportal		51	
Perisinusoidal and portal/periportal		32	
Bridging fibrosis		34	
Cirrhosis		11	

ALT, alanine transaminase; AST, aspartate aminotransferase; γ-GTP, gamma-glutamyl transferase.

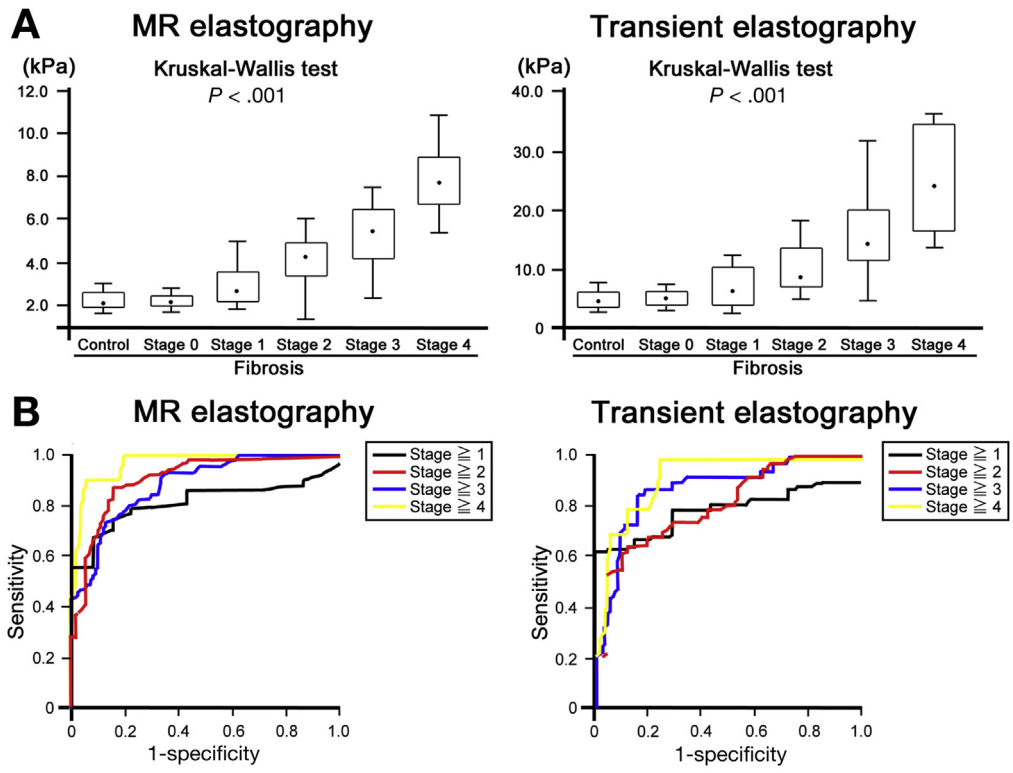


Figure 1. Relationship between LSM obtained using MRE or TE and liver fibrosis stage in control subjects and patients with NAFLD. (*A*) A steady stepwise increase in LSM for both MRE and TE was observed with increasing severity of liver fibrosis (Kruskal-Wallis test; $P < .001$). (*B*) The diagnostic accuracy of the LSM obtained using MRE or TE for liver fibrosis stage in patients with NAFLD. The AUROC curve is shown regarding the performance of the LSM in distinguishing liver fibrosis stage 0 from stages 1–4, 0–1 from 2–4, 0–2 from 3–4, and 0–3 from 4.

The purpose of this study was to compare clinical scoring systems, TE and MRE for the staging of liver fibrosis, and to compare TE-based CAP and MRI-based PDFF methods for grading hepatic steatosis in the same individuals.

Methods

Subject Characteristics

This cross-sectional study included patients evaluated at Yokohama City University Hospital, Yokohama, Japan. The cohort consisted of 142 patients, all of whom had liver biopsy–diagnosed NAFLD, including 34 with NAFL and 108 with NASH, and they were enrolled between July 2013 and April 2015. The time interval between clinical scoring systems, MRI, TE, and liver biopsy was <6 months. Patients with a history of excessive alcohol consumption (weekly consumption >140 g for men or >70 g for women), other liver diseases, such as chronic hepatitis, drug use associated with fatty liver, weight reduction, renal diseases, or thyroid disorders were excluded. The control group consisted of 10 subjects with a mean age and sex ratio comparable with those of the NAFL and NASH groups and all had normal liver enzyme levels; ultrasonography and liver biopsy showed no evidence of a fatty liver. The study protocol was reviewed and approved by the ethics review committee of each institution, and all subjects provided written informed consent before examination. The study protocol conformed to the ethical guidelines of the Declaration of Helsinki and was conducted with the approval of the Ethics Committee of Yokohama City University Hospital. This trial is registered with the UMIN Clinical Trials Registry as No. UMIN000012757.

Histopathologic and Immunohistochemical Evaluations

Liver biopsy samples were obtained from all patients with NAFLD using a 16-gauge needle biopsy kit according to a

Table 2. Diagnostic Accuracy of MRE and TE in Detecting Each Stage of Liver Fibrosis

	MRE (n = 142)							TE (n = 127)						
Fibrosis stage	Cut-off level, *kPa*	AUROC	95% CI	Se	Sp	PPV	NPV	Cut-off level, *kPa*	AUROC	95% CI	Se	Sp	PPV	NPV
≥1	2.5	0.80	0.71–0.89	75.0	85.7	99.0	84.6	7.0	0.78	0.70–0.87	61.7	100.0	100.0	86.6
≥2	3.4	0.89	0.85–0.94	87.3	85.0	88.4	83.6	11.0	0.82	0.74–0.89	65.2	88.7	88.2	66.2
≥3	4.8	0.89	0.83–0.95	74.5	86.9	74.5	81.0	11.4	0.88	0.79–0.97	85.7	83.8	75.0	91.9
≥4	6.7	0.97	0.94–1.00	90.9	94.5	58.8	99.2	14.0	0.92	0.86–0.98	100.0	75.9	73.0	100.0

NPV, negative predictive value; PPV, positive predictive value; Se, sensitivity; Sp, specificity.

Table 3. Diagnostic Accuracy of MRE, TE, and Scoring Systems in Detecting Each Stage of Liver Fibrosis

	Fibrosis stage (n = 127)															
	Stage 0 vs stage 1–4				Stage 0–1 vs stage 2–4				Stage 0–2 vs stage 3–4				Stage 0–3 vs stage 4			
Modality	AUROC	95% CI	*P* value	vs MRE *P* value	AUROC	95% CI	*P* value	vs MRE *P* value	AUROC	95% CI	*P* value	vs MRE *P* value	AUROC	95% CI	*P* value	vs MRE *P* value
MRE	0.83	0.72–0.93	.003		0.91	0.86–0.96	<.001		0.89	0.83–0.94	<.001		0.97	0.94–1.00	<.001	
TE	0.78	0.70–0.87	.003	.466	0.82	0.74–0.89	<.001	.001[a]	0.88	0.79–0.97	<.001	.426	0.92	0.86–0.98	<.001	.049[a]
FIB-4 index[b]	0.80	0.68–0.93	.003	.712	0.83	0.76–0.90	<.001	.023[a]	0.86	0.79–0.92	<.001	.384	0.88	0.82–0.95	.001	.036[a]
NFS[c]	0.82	0.70–0.93	.002	.870	0.82	0.74–0.89	<.001	.006[a]	0.860	0.80–0.92	<.001	.435	0.92	0.86–0.97	<.001	.132
APRI[d]	0.61	0.35–0.86	.443	.118	0.54	0.44–0.64	.402	<.001[a]	0.61	0.51–0.71	.130	<.001[a]	0.65	0.46–0.83	.128	.002[a]
AAR[e]	0.62	0.42–0.82	.155	.089	0.77	0.69–0.85	<.001	.003[a]	0.72	0.63–0.81	.002	.001[a]	0.71	0.56–0.86	.143	<.001[a]
BARD[f]	0.65	0.43–0.87	.164	.172	0.73	0.65–0.82	.002	<.001[a]	0.693	0.66–0.73	.002	.002[a]	0.68	0.57–0.80	.019	<.001[a]

[a]MRE was significantly superior to the modality.
[b]FIB-4 index = age (y) × aspartate aminotransferase (AST) (IU/L)/[platelets (10^9/L) × alanine transaminase (ALT)1/2 (IU/L)].
[c]NFS = −1.675 + 0.037 × age (y) + 0.094 × BMI (kg/m^2) + 1.13 × impaired fasting glycemia/diabetes (yes = 1, no = 0) + 0.99 × AST/ALT ratio − 0.013 × platelet (×10^9/L) − 0.66 × albumin (g/dL).
[d]APRI = [(AST/upper limit of normal AST) × 100]/platelets (10^9/L).
[e]AAR = AST/ALT ratio.
[f]BARD score = AAR ≥0.8 (2 points), BMI ≥28 (1 point), presence of diabetes (1 point).

standard protocol; 2 specimens were obtained from each patient to acquire a sample of sufficient size for analysis and to reduce histologic errors. An adequate liver biopsy sample was defined as being >20 mm in length and/or with >10 portal tracts. Livers were assessed histologically by 2 pathologists. Macrovesicular steatosis affecting at least 5% of hepatocytes was observed in all patients with NAFLD; these patients were classified as having (NASH) or not having steatohepatitis (NAFL). Patients with steatosis, inflammation, ballooned hepatocytes, and pericellular/perisinusoidal fibrosis were classified as having NASH.[30] Steatosis, lobular inflammation, and ballooned hepatocytes were classified as shown in the Supplementary Material. Fibrosis severity was scored as described previously.[31] Patients with NASH-associated cirrhosis were defined clinicopathologically.[32]

Magnetic Resonance Elastography

All of the included patients underwent hepatic MRE examinations performed using 3.0-T imagers (GE Healthcare, Milwaukee, WI) located at several different clinical sites on our campus; a 2-dimensional MRE protocol was used, similar to one described previously in the literature between December 2013 and April 2015.[23] MRE methods are detailed in the Supplementary Material. Interpretation of MRE images was performed by abdominal radiologists in the Department of Radiology following protocols established in the department, as described previously.[25] The LSM obtained at the time of examination was entered into the database and extracted for the present study.

Assessment of Steatosis Using Magnetic Resonance Imaging–Based Proton Density Fat Fraction

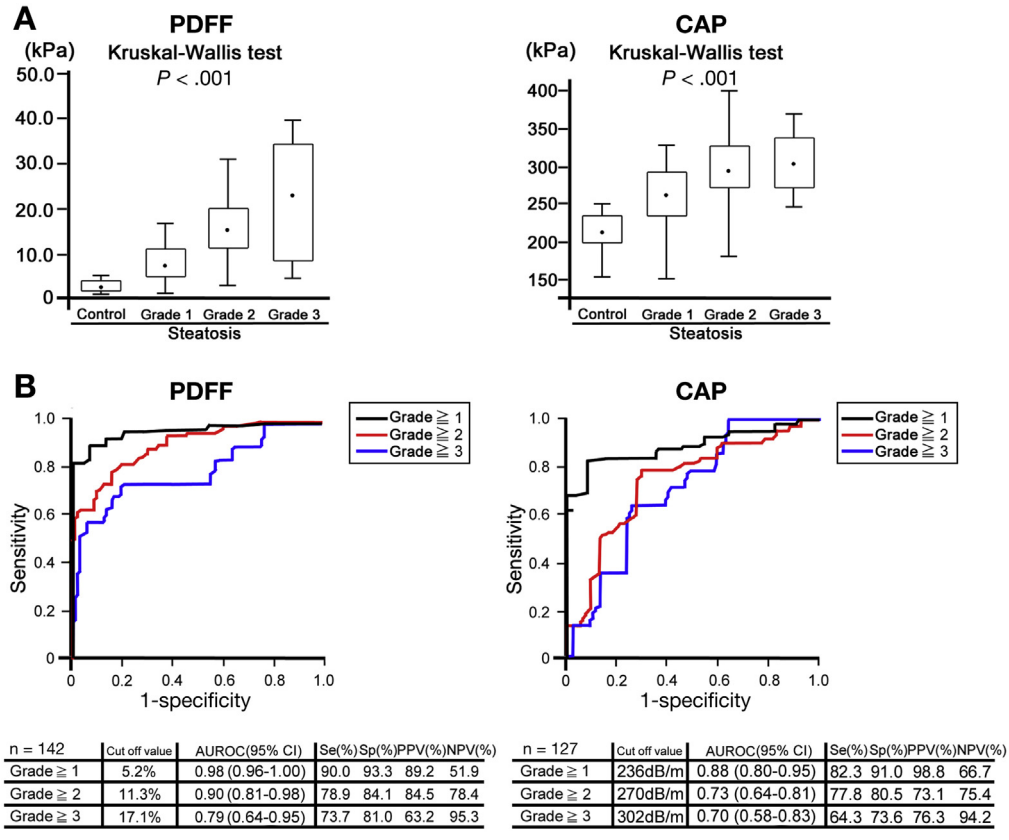
PDFF was measured using a modified Dixon method with advanced processing (IDEAL IQ, GE Healthcare).[33–35] PDFF methods are described in the Supplementary Material. Close to the regions of interest (ROIs) drawn for the LSM, new ROIs were drawn on the in-phase and out-of phase images for PDFF measurements. The PDFF was calculated as reported previously.[33–35] LSM and PDFF were analyzed by one author who was blinded to the liver histology results.

Transient Elastography

The LSM was obtained using a TE (M-probe, Fibroscan: EchoSens, Paris, France) by one operator. Details of the technique and the examination procedure have been described in previous reports.[10,36] Additionally, TE methods are detailed in the Supplementary Material. At the same time, hepatic steatosis was assessed using the CAP value provided by the device, only when the LSM was valid for the same signals, ensuring that the liver ultrasonic attenuation was obtained simultaneously from the same volume of liver parenchyma as the LSM.

Scoring Systems

Based on a review of the literature, the following scores were calculated for each patient: Fibrosis-4 index[37]; NAFLD fibrosis score[38]; aspartate aminotransferase to platelet ratio index,[39] aspartate aminotransferase to alanine transaminase ratio, and the BARD score.[40]



n = 142	Cut off value	AUROC(95% CI)	Se(%)	Sp(%)	PPV(%)	NPV(%)
Grade ≧ 1	5.2%	0.98 (0.96-1.00)	90.0	93.3	89.2	51.9
Grade ≧ 2	11.3%	0.90 (0.81-0.98)	78.9	84.1	84.5	78.4
Grade ≧ 3	17.1%	0.79 (0.64-0.95)	73.7	81.0	63.2	95.3

n = 127	Cut off value	AUROC(95% CI)	Se(%)	Sp(%)	PPV(%)	NPV(%)
Grade ≧ 1	236dB/m	0.88 (0.80-0.95)	82.3	91.0	98.8	66.7
Grade ≧ 2	270dB/m	0.73 (0.64-0.81)	77.8	80.5	73.1	75.4
Grade ≧ 3	302dB/m	0.70 (0.58-0.83)	64.3	73.6	76.3	94.2

Figure 2. Relationship between steatosis assessments obtained using MRI-based PDFF and TE-based CAP methods and pathology-based steatosis grade in control subjects and patients with NAFLD. (*A*) A steady stepwise increase in both PDFF and CAP was observed with increasing grade of steatosis (Kruskal-Wallis test; $P < .001$). (*B*) The diagnostic accuracy of the PDFF and CAP methods in assessing steatosis in control subjects and patients with NAFLD. The AUROC curve is shown for the performance of the PDFF or CAP in distinguishing steatosis grade 0 (controls) from grades 1–3, 0–1 from 2–3, and 0–2 from 3.

Table 4. Diagnostic Accuracy of PDFF and CAP in Detecting Each Grade of Steatosis

Steatosis grade	PDFF (n = 127)							CAP (n = 127)							
	Cutoff level, %	AUROC	95% CI	Se, %	Sp, %	PPV, %	NPV, %	Cutoff level, *dB/m*	AUROC	95% CI	Se, %	Sp, %	PPV, %	NPV, %	vs PDFF *P* value
≥1	5.2	0.96	0.92–1.00	90.0	93.3	89.2	51.9	236	0.88	0.80–0.95	82.3	91.0	98.9	66.7	.048[a]
≥2	11.3	0.90	0.82–0.97	78.9	84.1	84.5	78.4	270	0.73	0.64–0.81	77.8	80.5	73.1	75.4	<.001[a]
≥3	17.1	0.79	0.65–0.94	73.7	81.0	63.2	95.3	302	0.70	0.58–0.83	64.3	73.6	76.3	94.2	.015[a]

NPV, negative predictive value; PPV, positive predictive value; Se, sensitivity; Sp, specificity.
[a]PDFF was significantly superior to CAP.

Statistical Analysis

Continuous variables were summarized as means and SDs and categorical variables as frequencies and percentages. All statistical analyses were performed using SPSS software (version 12, SPSS Inc., Chicago, IL). The *t* test and analysis of variance with Scheffé multiple testing correction were used for univariate comparisons between groups. Because many of the variables were not normally distributed, the Kruskal-Wallis test was used for comparisons of >2 independent groups. The *z* test was used for comparisons of area under the receiver operating characteristic (AUROC) between 2 groups.[41] *P* values <.05 were considered statistically significant.

All authors had access to the study data and reviewed and approved the final manuscript.

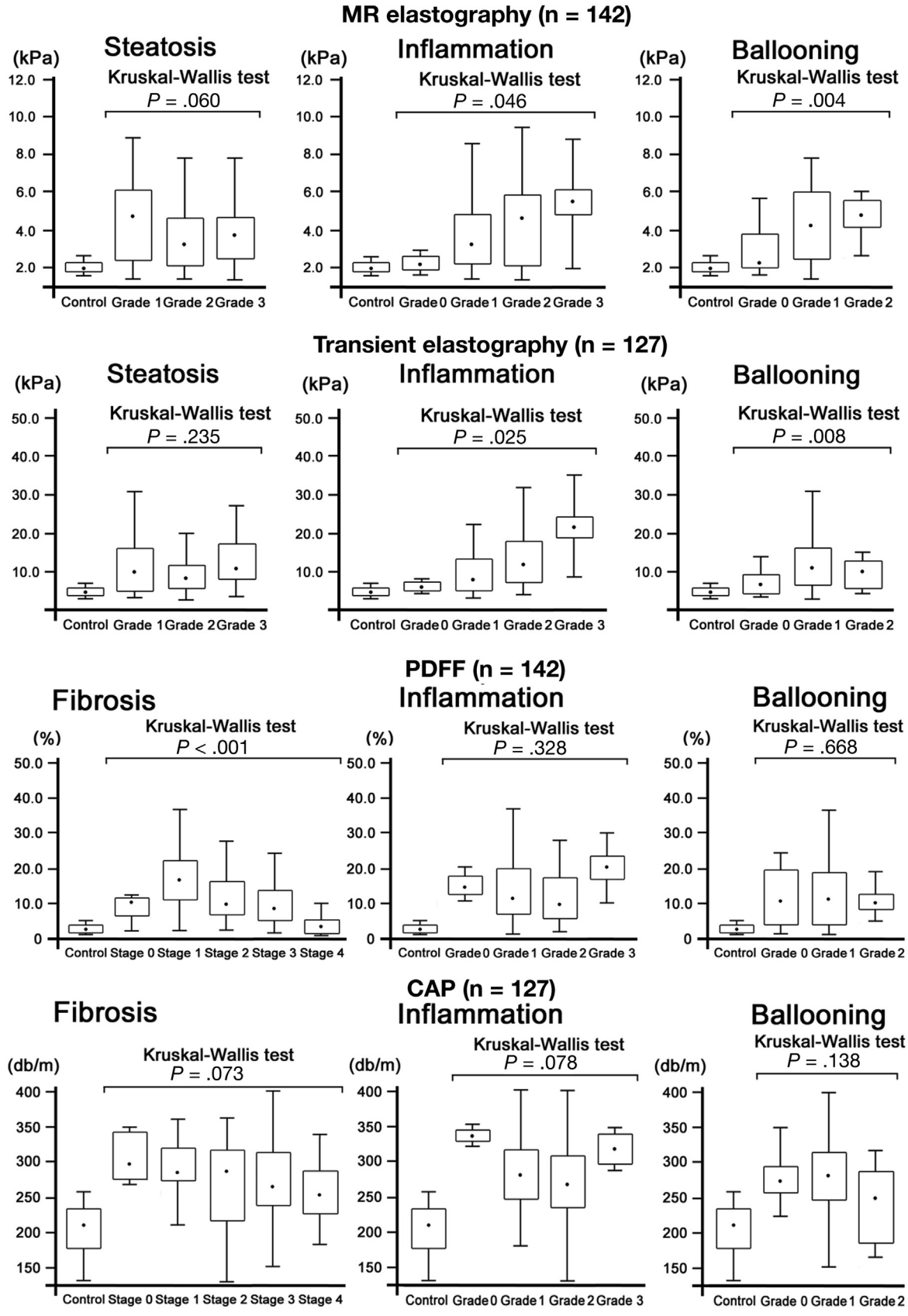
Results

Patient Characteristics

In this cross-sectional study, there were 10 control subjects and a total of 142 patients with NAFLD who underwent MRE and PDFF measurements using MRI. TE and CAP assessment was attempted in all of the control subjects and patients, but 15 of these examinations failed because of unreliable LSM and CAP (no successful acquisitions, Supplementary Table 1). Table 1 details the principal features and laboratory characteristics of the enrolled subjects. Histologic characteristics are also summarized in Table 1.

Assessment of Liver Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Using Clinical Scoring Systems, Transient Elastography, and Magnetic Resonance Elastography

The LSM was measured using MRE and TE in patients with NAFLD to assess the stage of liver fibrosis. The mean LSM values (shear modulus-based) for MRE (n = 142) were 2.11, 2.16, 2.62, 4.28, 5.24, and 7.93 kPa for the controls and stages 0, 1, 2, 3, and 4, respectively. The mean LSM values (Young's modulus-based) for TE (n = 127), in patients where it was measured successfully, were 5.32, 5.50, 7.61, 9.42, 14.70, and 24.23 kPa for the controls and stages 0, 1, 2, 3, and 4, respectively, as shown in Figure 1*A* and *B*. The results of these analyses revealed stepwise increases in the LSM obtained using MRE or TE with increasing histologic severity of hepatic fibrosis ($P < .001$ using the Kruskal-Wallis test). To investigate the diagnostic accuracy of the LSM obtained using MRE or TE for liver fibrosis, we calculated the AUROC curve and potential cutoff values for the diagnosis of liver fibrosis in NAFLD patients. The ROC curves for differentiating between liver fibrosis stage 0 and stages 1–4, 0–1 and 2–4, 0–2 and 3–4, and 0–3 and 4 based on the LSM measured using MRE or TE in NAFLD patients are shown in Figure 1*B*. The AUROC curve in diagnosing liver fibrosis stage ≥1, ≥2, ≥3, and ≥4 using MRE or TE were 0.80 (95% confidence interval [CI]: 0.71–0.90) or 0.78 (95% CI: 0.70–0.87), 0.89 (95% CI: 0.85–0.94) or 0.82 (95% CI: 0.74–0.90), 0.89 (95% CI: 0.83–0.95) or 0.88 (95% CI: 0.79–0.97), and 0.97 (95% CI: 0.94–1.00) or 0.92 (95% CI: 0.86–0.98), respectively. The cutoff level and the sensitivity, specificity, positive predictive


MR elastography (n = 142)
Steatosis
(kPa)
Kruskal-Wallis test
P = .060
Control Grade 1 Grade 2 Grade 3
Inflammation
(kPa)
Kruskal-Wallis test
P = .046
Control Grade 0 Grade 1 Grade 2 Grade 3
Ballooning
(kPa)
Kruskal-Wallis test
P = .004
Control Grade 0 Grade 1 Grade 2
Transient elastography (n = 127)
Steatosis
(kPa)
Kruskal-Wallis test
P = .235
Control Grade 1 Grade 2 Grade 3
Inflammation
(kPa)
Kruskal-Wallis test
P = .025
Control Grade 0 Grade 1 Grade 2 Grade 3
Ballooning
(kPa)
Kruskal-Wallis test
P = .008
Control Grade 0 Grade 1 Grade 2
PDFF (n = 142)
Fibrosis
(%)
Kruskal-Wallis test
P < .001
Control Stage 0 Stage 1 Stage 2 Stage 3 Stage 4
Inflammation
(%)
Kruskal-Wallis test
P = .328
Control Grade 0 Grade 1 Grade 2 Grade 3
Ballooning
(%)
Kruskal-Wallis test
P = .668
Control Grade 0 Grade 1 Grade 2
CAP (n = 127)
Fibrosis
(db/m)
Kruskal-Wallis test
P = .073
Control Stage 0 Stage 1 Stage 2 Stage 3 Stage 4
Inflammation
(db/m)
Kruskal-Wallis test
P = .078
Control Grade 0 Grade 1 Grade 2 Grade 3
Ballooning
(db/m)
Kruskal-Wallis test
P = .138
Control Grade 0 Grade 1 Grade 2

value, and negative predictive value for each liver fibrosis stage are detailed in Table 2. The results indicate that MRE had high diagnostic accuracy in the assessment of liver fibrosis compared with TE. In Table 3, data present direct comparisons regarding the diagnostic accuracy of noninvasive markers, including LSM and scoring systems (such as the Fibrosis-4 index, NAFLD fibrosis score, aspartate aminotransferase to platelet ratio index, aspartate aminotransferase to alanine transaminase ratio, and BARD score) in detecting liver fibrosis in patients with NAFLD (n = 127), in which both MRE and TE could be performed successfully. There were significantly greater diagnostic accuracy for liver fibrosis stage ≥2 and ≥4 in MRE than that in TE (stage ≥2, MRE: AUROC = 0.91; 95% CI: 0.86−0.96 vs TE: AUROC = 0.82; 95% CI: 0.74−0.89; $P = .001$; stage 4 or higher, MRE: AUROC = 0.97; 95% CI: 0.94−1.00 vs TE: AUROC: 0.92; 95% CI: 0.86−0.98; $P = .049$). The AUROC curve results indicate that MRE had high diagnostic accuracy in the assessment of each liver fibrosis stage relative to TE and all clinical scoring systems, and was significantly superior to TE and the clinical scoring systems in the diagnosis of the fibrosis stage (Table 3).

Assessment of Steatosis in Patients With Nonalcoholic Fatty Liver Disease Using Magnetic Resonance Imaging−Based Proton Density Fat Fraction and Transient Elastography−Based Controlled Attenuation Parameter Methods

PDFF measurements obtained using MRI, and CAP-based measurements obtained using TE, were compared with the steatosis grade obtained by liver biopsy. The mean PDFF measurements were 3.2% for control subjects (grade 0), 8.1% for grade 1, 16.3% for grade 2, and 22.9% for grade 3 (Figure 2*A*). In contrast, the CAP value was 210.7 dB/m for control subjects (grade 0), 262.9 dB/m for grade 1, 289.6 dB/m for grade 2, and 304.9 dB/m for grade 3 (Figure 2*A*). ROC curves for the PDFF (n = 142) and CAP (n = 127), in patients where it was successfully measured, were plotted for steatosis grade, including the control subjects. For the detection of grade ≥1, the AUROC curve for PDFF and CAP was 0.98 (95% CI: 0.96−1.00) and 0.88 (95% CI: 0.80−0.95), respectively, using ROC analysis (see grade ≥1 in Figure 2*B*). Next, for the detection of grade ≥2 as compared with grade ≤1, the AUROC curve for the PDFF and CAP was 0.90 (95% CI: 0.81−0.98) and 0.73 (95% CI: 0.64−0.81), respectively (see grade ≥2 in Figure 2*B*). Finally, for the detection of grade ≥3 as compared with grade ≤2, the AUROC curve for the PDFF and CAP was 0.79 (95% CI: 0.64−0.95) and 0.70 (95% CI: 0.58−0.83), respectively (see grade ≥3 in Figure 2*B*). The sensitivity, specificity, positive predictive value, and negative predictive value in the diagnosis of steatosis grade in NAFLD were calculated based on the ROC analyses (Figure 2*B*). Additionally, the diagnostic accuracy of PDFF and CAP in detecting steatosis in patients with NAFLD (n = 127) in which both PDFF and CAP could be performed successfully, are directly compared in Table 4. There were significantly greater diagnostic accuracy for steatosis grade PDFF than that in CAP (grade ≥1 PDFF:, AUROC = 0.96; 95% CI: 0.92−1.00; vs CAP: AUROC = 0.88; 95% CI: 0.80−0.95; $P = .048$. Grade 2 or higher PDFF: AUROC = 0.90; 95% CI: 0.82−0.97 vs CAP: AUROC = 0.73; 95% CI: 0.64−0.81; $P < .001$; grade 3 or higher PDFF: AUROC = 0.79; 95% CI: 0.65−0.94 vs CAP: AUROC = 0.70; 95% CI: 0.58−0.83; $P = .015$). The AUROC curve results indicate that PDFF was significantly superior to CAP in the diagnosis of each steatosis grade.

Correlation Between Liver Stiffness Measurements Based on Magnetic Resonance Elastography or Transient Elastography, and Steatosis Measurements Based on Proton Density Fat Fraction or Controlled Attenuation Parameter and Histologic Parameters

We investigated the relationship between degree of steatosis, hepatic inflammatory activity, or ballooned hepatocytes and LSM obtained using MRE or TE, and between the degrees of liver fibrosis, hepatic inflammatory activity, or ballooned hepatocytes and PDFF and CAP-based steatosis measurements in NAFLD patients. There was a significant correlation between LSM and grade of hepatic inflammatory activity and ballooned hepatocytes in both MRE and TE (Figure 3*A*), and between PDFF and the stage of liver-only fibrosis, but not with the CAP-based measurements (Figure 3*B*).

Multiple Regression Analysis of Histologic Parameters Associated With Liver Stiffness Measurement Based on Magnetic Resonance Elastography or Transient Elastography, and Steatosis Measurements Based on Proton Density Fat Fraction or Controlled Attenuation Parameter

Next, the relationship between histologic parameters and LSM obtained using MRE or TE, and PDFF or CAP measurements was studied using multivariate regression analysis (Supplementary Table 2). Only liver fibrosis stage was significantly correlated with LSM using multiple regression analysis. In addition, liver fibrosis stage and steatosis grade or only steatosis grade were correlated significantly with the MRI-based PDFF measurements or TE-based CAP measurements using multiple regression analysis.

Figure 3. Correlation between the LSM obtained using MRE or TE, and steatosis measurements obtained using PDFF and CAP methods and histologic parameters in control subjects and patients with NAFLD. There was a significant correlation between LSM and the grade of hepatic inflammatory activity and ballooning degeneration regarding both MRE and TE in patients with NAFLD. The vertical axis represents the LSM in kPa (shear modulus for MRE and Young's modulus for TE) and the horizontal axis represents each grade of steatosis, inflammation and ballooned hepatocyte. There was significant correlation between PDFF and liver fibrosis stage in patients with NAFLD, but not CAP.

Assessment of Nonalcoholic Steatohepatitis or Nonalcoholic Fatty Liver Disease Activity Score ≥5 in Patients With Nonalcoholic Fatty Liver Disease Using Magnetic Resonance Imaging and Transient Elastography Methods With Serum Keratin 18 and Alanine Transaminase Levels

We assessed the diagnostic ability of the LSM obtained using MRE combined with MRI-based PDFF (MRE+PDFF) and the LSM obtained using TE including CAP (TE+CAP) regarding NASH and Nonalcoholic Fatty Liver Disease Activity Score (NAS) ≥5. We report the results in the Supplementary Material. The AUROC curve results indicate that MRE+PDFF was significantly superior to TE+CAP in the diagnosis of NAS ≥5 (Supplementary Table 3; MRE+PDFF: AUROC = 0.77; 95% CI: 0.67−0.87 vs TE+CAP: AUROC = 0.65; 95% CI: 0.54−0.77; $P = .045$.). Additionally, we assessed the diagnostic ability of MRE+PDFF or TE+CAP with serum keratin 18 or alanine transaminase levels concerning NASH and NAS ≥5. We report the results in the Supplementary Material. We found that the addition of serum keratin 18 and alanine transaminase levels did not improve the diagnostic ability of MRE+PDFF and TE+CAP significantly concerning NASH and NAS ≥5 (Supplementary Figure 1 and Supplementary Tables 4–6).

Discussion

This cross-sectional study demonstrated a significant positive correlation between the LSM obtained using MRE and the severity of liver fibrosis in patients with NAFLD. In addition, the diagnostic accuracy of MRE for liver fibrosis was found to be higher than that of clinical scoring systems and TE. Although a steady stepwise increase in PDFF and CAP was observed with increasing severity of hepatic steatosis, the diagnostic accuracy of PDFF regarding hepatic steatosis grade was also significantly superior to that of CAP. To our knowledge, this is the first study that provides a head-to-head comparison of LSM measured using TE and MRE, fat accumulation evaluated using CAP measured by means of TE, and PDFF measured using MRI for liver fibrosis and steatosis in biopsy-proven NAFLD.

NAFLD, which can progress to cirrhosis and hepatocellular carcinoma, is considered to be the most common form of chronic liver disease in obese patients. However, because effective therapies for NAFLD have not yet been established, the identification of risk factors for hepatocellular carcinoma, such as liver fibrosis, would help to guide the implementation of risk-reduction strategies for these patients.[42] Liver biopsy has been considered the reference standard in the assessment of liver fibrosis and steatosis, although its limitations are well recognized and include risk of complications and costs. In addition, the accuracy of the procedure used to assess the severity of liver fibrosis and steatosis is questionable because of intra- and inter-observer variation.[43,44] Sampling error has also been reported, even in patients with NASH.[9] Additionally, because NAFLD occurs in 25% of all adults in many countries, it is impossible to perform liver biopsy in all NAFLD patients. Therefore, alternatives to liver biopsy have been investigated, such as clinical scoring systems, TE, and MRI, which can be used repeatedly because of high safety. Although several scoring systems for the diagnosis for liver fibrosis in patients with NAFLD have been reported previously, the Fibrosis-4 index and NAFLD fibrosis score exhibited good performance characteristics in assessing fibrosis.[37,38] However, these scoring systems have limitations, and the positive predictive value is very low. Previous studies have indicated that TE can be used to measure the severity of liver fibrosis in patients with NAFLD.[11–13] Although TE is very useful in diagnosing the severity of liver fibrosis in NASH, its success rate is dependent on operator expertise, as well as on other factors (age, width of the intercostal space, ascites, BMI, and visceral fat); Sporea et al[45] have reported a rate of reliable measurements of 81.6%, which is in line with that of Castéra et al.[46] This failure was independently associated with a BMI >28 kg/m^2 because of the low-frequency vibrations induced by the probe.[18] Other authors have suggested that a high BMI or large amount of ascites cause the failure of TE.[47–49] Indeed, in our study, TE was unsuccessful in assessing LSM and CAP in 15 patients (about 10% of the total cohort; Supplementary Table 1). Although a new XL probe equipped with CAP, which became available from May 2015, has been reported to reduce scan failure and enhance the reliability of the measurement of fibrosis and steatosis in morbidly obese patients (BMI >28 kg/m^2),[47,49] the probe cannot be used in many countries, including Japan. In addition, the diagnostic ability of the XL probe seems to be similar to that of the M-probe in patients who are not morbidly obese (BMI ≤28 kg/m^2). In contrast, because MRE uses compressional and continuous waves, the elastic waves generated by the vibrator could be effective in patients with ascites or obesity, including patients with a BMI >28 kg/m^2.[50] Additionally, MRE offers more advantages than TE as follows: First, the 2-dimensional displacement vector is assessed in MRE, whereas only 1 directional measurement is performed in TE, which might be vulnerable to complex waves, including reflection and refraction.[35] Second, the area measured in the liver is larger in MRE than in TE, which can avoid the sampling variability caused by the heterogeneity of advanced fibrosis.[51,52] In fact, in our study, 9 (7.1%) patients with advanced fibrosis (stage 3−4) were diagnosed by TE as "mild fibrosis" out of 127 patients with NAFLD. In contrast, only 4 (2.8%) patients with advanced fibrosis were diagnosed using MRE as "mild fibrosis" out of 142 patients with NAFLD; this suggested that TE may misclassify patients with advanced fibrosis as compared with MRE. In contrast, MRE may misclassify patients with mild fibrosis (stage 0−2) relative to TE. Indeed, 12 (7.1%) patients with mild fibrosis were diagnosed using TE as having advanced fibrosis out of 127 patients with NAFLD. In contrast, 22 (15.5%) patients with mild fibrosis were diagnosed using MRE as having advanced fibrosis out of 142 patients with NAFLD. We consider that this misclassification using MRE was associated with sampling errors involved in the use of liver biopsy and TE.

Hepatic fat accumulation is receiving increasing attention in clinical practice because the prevalence of steatosis

associated with obesity is dramatically affecting developed countries.[53,54] In addition, it is also important for liver transplantation because a ≥30% fat content contraindicates liver donation.[55] Consequently, the establishment of new noninvasive approaches is essential in accurately determining the hepatic fat concentration, and in facilitating the correct diagnosis and monitoring of steatosis. To date, there have been no specific biochemical or serologic tests that are able to diagnose the presence of steatosis, or to quantitate the degree of steatosis. In our study, PDFF and CAP each correlated reasonably well with biopsy-proven steatosis, and significantly differentiated controls from NAFLD patients with steatosis grades of 1, 2, and 3. However, the diagnostic accuracy of MRI-based PDFF for each steatosis grade was superior to TE-based CAP in the control subjects and patients with NAFLD, in particular in the case of advanced steatosis (grades 2−3). Although MRE and PDFF seem to be superior to TE including CAP in the diagnosis of fibrosis and steatosis in patients with NAFLD, costs associated with the use of MRI are higher than those for TE (Supplementary Table 7). However, MRI can be used to scan the whole liver for screening with the objective of detecting hepatocellular carcinoma at the same time as measuring the LSM and PDFF. Therefore, considered overall, the cost benefit of MRI seems to be comparable with TE.

Our study had several limitations. First, the use of liver biopsy as the gold standard for assessing liver pathology has limitations associated with sampling errors, as well as intra- and inter-observer variability, which are at least partly linked to the size of the biopsy. Second, we did not consider the dynamic effect of hepatic perfusion in our study, which could cause elevated LSM unrelated to the liver disease itself. This effect has been observed in TE and MRE studies.[56,57] Third, the drawing of ROIs on the MRE image is not yet fully standardized. We measured the mean of 3 ROIs to decrease measurement bias in this study. Finally, patient selection bias may also have been the result of liver biopsies being more likely to be performed on NAFLD patients at risk for NASH with advanced fibrosis.

In conclusion, this is the first study that has demonstrated that MRI-based MRE and PDFF methods have higher diagnostic accuracy in detecting liver fibrosis and steatosis, respectively, in patients with NAFLD, relative to the TE-based LSM and CAP methods. This was confirmed by the liver biopsy results, which remains the gold standard for the evaluation of the severity of liver fibrosis and steatosis. Further studies must be conducted to explore the prognostic value of the results of these diagnostic techniques to determine the long-term outcomes of patients with NAFLD.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2015.11.048.

References

1. Lazo M, Clark JM. The epidemiology of nonalcoholic fatty liver disease: a global perspective. Semin Liver Dis 2008;28:339–350.
2. Day CP. Non-alcoholic steatohepatitis (NASH): where are we now and where are we going? Gut 2002;50:585–588.
3. Hamaguchi M, Kojima T, Takeda N, et al. The metabolic syndrome as a predictor of nonalcoholic fatty liver disease. Ann Intern Med 2005;143:722–728.
4. Matteoni C, Younossi ZM, Gramlich T, et al. A non-alcoholic fatty liver disease: a spectrum of clinical and pathologic severity. Gastroenterology 1999;116:1413–1419.
5. Day CP, Saksena S. Non-alcoholic steatohepatitis: definitions and pathogenesis. J Gastroenterol Hepatol 2002;17:377–384.
6. Harrison SA, Torgerson S, Hayashi PH, et al. The natural history of nonalcoholic fatty liver disease: a clinical histopathological study. Am J Gastroenterol 2003;98:2042–2047.
7. Angulo P. Nonalcoholic fatty liver disease. N Engl J Med 2002;18:1221–1231.
8. Cadranel JF. Good clinical practice guidelines for fine needle aspiration biopsy of the liver: past, present and future. Gastroenterol Clin Biol 2002;26:823–824.
9. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898–1906.
10. Sandrin L, Tanter M, Gennisson JL, et al. Shear elasticity probe for soft tissues with 1-D transient elastography. IEEE Trans Ultrason Ferroelectr Freq Control 2002;49:436–446.
11. Yoneda M, Yoneda M, Fujita K, et al. Transient elastography in patients with non-alcoholic fatty liver disease (NAFLD). Gut 2007;56:1330–1331.
12. Yoneda M, Yoneda M, Mawatari H, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with nonalcoholic fatty liver disease (NAFLD). Dig Liver Dis 2008;40:371–378.
13. Kwok R, Tse YK, Wong GL, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease—the role of transient elastography and plasma cytokeratin-18 fragments. Aliment Pharmacol Ther 2014;39:254–269.
14. Sasso M, Beaugrand M, de Ledinghen V, et al. Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. Ultrasound Med Biol 2010;36:1825–1835.
15. Alisi A, Pinzani M, Nobili V. Diagnostic power of fibroscan in predicting liver fibrosis in nonalcoholic fatty liver disease. Hepatology 2009;50:2048–2049.
16. Chan WK, Nik Mustapha NR, Mahadeva S. Controlled attenuation parameter for the detection and quantification of hepatic steatosis in nonalcoholic fatty liver disease. J Gastroenterol Hepatol 2014;29:1470–1476.
17. de Lédinghen V, Vergniol J, Capdepont M, et al. Controlled attenuation parameter (CAP) for the diagnosis of steatosis: a prospective study of 5323 examinations. J Hepatol 2014;60:1026–1031.

18. Foucher J, Castéra L, Bernard PH, et al. Prevalence and factors associated with failure of liver stiffness measurement using FibroScan in a prospective study of 2114 examinations. Eur J Gastroenterol Hepatol 2006; 18:411–412.
19. Carrión JA, Navasa M, Forns X. MR elastography to assess liver fibrosis. Radiology 2008;247:591; author reply 591–592.
20. Talwalkar JA. Elastography for detecting hepatic fibrosis: options and considerations. Gastroenterology 2008; 135:299–302.
21. Talwalkar JA, Yin M, Fidler JL, et al. Magnetic resonance imaging of hepatic fibrosis: emerging clinical applications. Hepatology 2008;47:332–342.
22. Huwart L, Sempoux C, Vicaut E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology 2008;135:32–40.
23. Yin M, Talwalkar JA, Glaser KJ, et al. Assessment of hepatic fibrosis with magnetic resonance elastography. Clin Gastroenterol Hepatol 2007;5:1207–1213.
24. Muthupillai R, Lomas DJ, Rossman PJ, et al. Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. Science 1995; 269(5232):1854–1857.
25. Chen J, Talwalkar JA, Yin M, et al. Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. Radiology 2011;259:749–756.
26. Kim D, Kim WR, Talwalkar JA, et al. Advanced fibrosis in nonalcoholic fatty liver disease: noninvasive assessment with MR elastography. Radiology 2013;268:411–419.
27. **Loomba R, Sirlin C**, et al. Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: a prospective study. Hepatology 2014;60:1920–1928.
28. Reeder SB, Robson PM, Yu H, et al. Quantification of hepatic steatosis with MRI: the effects of accurate fat spectral modeling. J Magn Reson Imaging 2009; 29:1332–1339.
29. Permutt Z, Le TA, Peterson MR, et al. Correlation between liver histology and novel magnetic resonance imaging in adult patients with non-alcoholic fatty liver disease-MRI accurately quantifies hepatic steatosis in NAFLD. Aliment Pharmacol Ther 2012;36:22–29.
30. Younossi ZM, Stepanova M, Rafiq N, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. Hepatology 2011;53:1874–1882.
31. Brunt EM. Nonalcoholic steatohepatitis: definition and pathology. Semin Liver Dis 2001;21:3–16.
32. Hui JM, Kench JG, Chitturi S, et al. Long-term outcomes of cirrhosis in nonalcoholic steatohepatitis compared with hepatitis C. Hepatology 2003;38:420–427.
33. Levenson H, Greensite F, Hoefs J, et al. Fatty infiltration of the liver: quantification with phase-contrast MR imaging at 1.5 T vs biopsy. AJR Am J Roentgenol 1991; 156:307–312.
34. Bernstein MA, King KF, Zhou XJ. Handbook of MRI Pulse Sequences. San Diego, CA: Elsevier Academic Press, 2004.
35. Hussain HK, Chenevert TL, Londy FJ, et al. Hepatic fat fraction: MR imaging for quantitative measurement and display-early experience. Radiology 2005;237:1048–1055.
36. Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 2003;29:1705–1713.
37. Sumida Y, Yoneda M, Hyogo H, et al; Japan Study Group of Nonalcoholic Fatty Liver Disease (JSG-NAFLD). Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. BMC Gastroenterol 2012;12:2.
38. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 2007; 45:846–854.
39. Wai CT, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003;38:518–526.
40. Harrison SA, Oliver D, Arnold HL, et al. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut 2008; 57:1441–1447.
41. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845.
42. Sanyal AJ. Treatment of non-alcoholic fatty liver disease. J Gastroenterol Hepatol 2002;17:385–388.
43. Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol 2002; 97:2614–2618.
44. Bedossa P, Carrat F. Liver biopsy: the best, not the gold standard. J Hepatol 2009;50:1–3.
45. Sporea I, Jurchiş A, Şirli R, et al. Can transient elastography be a reliable method for assessing liver fibrosis in nonalcoholic steatohepatitis (NASH)? Med Ultrason 2013;15:106–110.
46. Castéra L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13, 369 examinations. Hepatology 2010;51:828–835.
47. de Lédinghen V, Vergniol J, Foucher J, et al. Feasibility of liver transient elastography with FibroScan using a new probe for obese patients. Liver Int 2010;30:1043–1048.
48. Chang PE, Lui HF, Chau YP, et al. Prospective evaluation of transient elastography for the diagnosis of hepatic fibrosis in Asians: comparison with liver biopsy and aspartate transaminase platelet ratio index. Aliment Pharmacol Ther 2008;28:51–61.
49. Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol 2012; 107:1862–1871.
50. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of the liver: technique, analysis, and clinical applications. J Magn Reson Imaging 2013;37:544–555.
51. Faria SC, Ganesan K, Mwangi I, et al. MR imaging of liver fibrosis: current state of the art. Radiographics 2009; 29:1615–1635.

52. Hines CD, Bley TA, Lindstrom MJ, et al. Repeatability of magnetic resonance elastography for quantification of hepatic stiffness. J Magn Reson Imaging 2010;31:725–731.
53. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 2011;34:274–285.
54. Adams LA, Lindor KD. Nonalcoholic fatty liver disease. Ann Epidemiol 2007;17:863–869.
55. Kim SH, Lee JM, Han JK, et al. Hepatic macrosteatosis: predicting appropriateness of liver donation by using MR imaging–correlation with histopathologic findings. Radiology 2006;240:116–129.
56. Yin M, Kolipaka A, Woodrum DA, et al. Hepatic and splenic stiffness augmentation assessed with MR elastography in an in vivo porcine portal hypertension model. J Magn Reson Imaging 2013;38:809–815.
57. Bureau C, Metivier S, Peron JM, et al. Transient elastography accurately predicts presence of significant portal hypertension in patients with chronic liver disease. Aliment Pharmacol Ther 2008;27:1261–1268.

Author names in bold designate shared co-first authorship.

Received May 31, 2015. Accepted November 21, 2015.

Reprint requests
Address requests for reprints to: Atsushi Nakajima, MD, PhD, Division of Gastroenterology, Yokohama City University Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan. e-mail: nakajima-tky@umin.ac.jp; fax: +81-45-784-3546.

Acknowledgments
The skillful technical assistance of Machiko Hiraga and Ayama Miura is gratefully acknowledged. All authors had access to the study data and reviewed and approved the final manuscript.

Conflicts of interest
The authors disclose no conflicts.

Funding
Work in the authors' laboratory was supported by the "Step A" program of the Japan Science and Technology Agency and Kiban-B, Shingakujuturyouiki. In addition, this work was supported in part by grants-in-aid from the Japanese Ministry of Health, Labour and Welfare.

Supplementary Methods

Histopathologic and Immunohistochemical Evaluations

One of the 2 pathologists was a hepatopathologist. Steatosis affecting <5%, 5%−33%, 33%−66%, and >66% of hepatocytes was classified as grades 0, 1, 2, and 3, respectively. Lobular inflammation was graded according to the number of inflammatory foci per field of view (FOV) at a magnification of 200×, with 0, <2, 2−4, and >4 foci per field classified as grades 0, 1, 2 and 3, respectively. Hepatocellular ballooning involving no, few, and many cells was classified as grades 0, 1, and 2, respectively.

Magnetic Resonance Elastography

Continuous longitudinal mechanical waves (60 Hz) were generated using a passive acoustic driver placed against the anterior chest wall. A 2-dimensional spin-echo planar MRE sequence was used to acquire axial wave images with the following parameters: repetition time/echo time: 50/23 ms; continuous sinusoidal vibration: 60 Hz; FOV: 32−42 cm; matrix size: 256 × 64; flip angle: 30 degrees; section thickness: 10 mm; 4 evenly spaced phase offsets; and 4 pairs of 60-Hz trapezoidal motion encoding gradients with zeroth and first moment nulling along the through-plane direction. All processing steps were applied automatically, without manual intervention, to yield quantitative images of tissue shear stiffness in kilopascals. On each section of the MR magnitude image from the MRE acquisition, the ROIs were drawn to include only the parenchyma of the liver, avoiding the edges of the liver and large blood vessels. The ROIs also excluded regions where the phase signal-to-noise ratio (the ratio of wave amplitude to the noise in the wave images) was <5.

Assessment of Steatosis Using Magnetic Resonance Imaging−Based Proton Density Fat Fraction

A fast gradient echo sequence was performed to acquire in-phase and out-of-phase images with the following parameters: repetition time/echo times: 110/2.1 ms; 4.3; flip angle: 70 degree; matrix: 256 × 192; axial imaging plane; section thickness: 6 mm; FOV: 34−44 cm; fractional phase FOV: 0.75−1; one signal acquired; bandwidth, 62.5 kHz; imaging time, 2 breath holds (about 16 seconds each).

Transient Elastography

TE was performed in control subjects and NAFLD patients under fasting conditions for 12 hours. This system is equipped with a probe and an ultrasonic transducer mounted on the axis of a vibrator. A vibration of mild amplitude and low frequency is transmitted from the vibrator to the tissue by the transducer itself; this induces the propagation of an elastic shear wave through the tissue. The speed of the propagating wave is estimated using a 1-dimensional ultrasound technique, and this is automatically converted to a measurement in terms of Young's modulus in units of kilopascals. TE-based measurements of Young's modulus can be compared approximately to MRE-based measurements of shear modulus by dividing the TE measurements by a factor of 3. In this study, only TE measurement with at least 10 valid shots and a success rate of ≥60% were considered reliable, and were used for statistical analysis.

Scoring Systems

The values for the upper limit of normal were set according to the International Federation of Clinical Chemistry: an aspartate aminotransferase level of 35 U/L for men and 30 U/L for women (these levels were comparable with those used in other analyses).

Keratin 18 fragments

Serum samples used in the current study were aliquots (0.5 mL) from the original samples in which blood from fasting subjects was collected into serum separator tubes, allowed to clot for at least 30 minutes at room temperature, and centrifuged at 1800*g* for 15 minutes. Aliquots of serum were immediately frozen. Processing was completed within 2 hours, and samples were free of hemolysis.

The serum caspase-cleaved K18 fragment (K18 [caspase-cleaved cytokeratin-18]) levels were determined using the M30-Apoptosense enzyme-linked immunosorbent assay (Peviva AB, Bromma, Sweden). Working range and detection limit of the assay kit is 75−1000 U/L and 20 U/L. The standard curve and K18 values were determined using the Soft-max Pro software accompanying the microplate reader.

Results

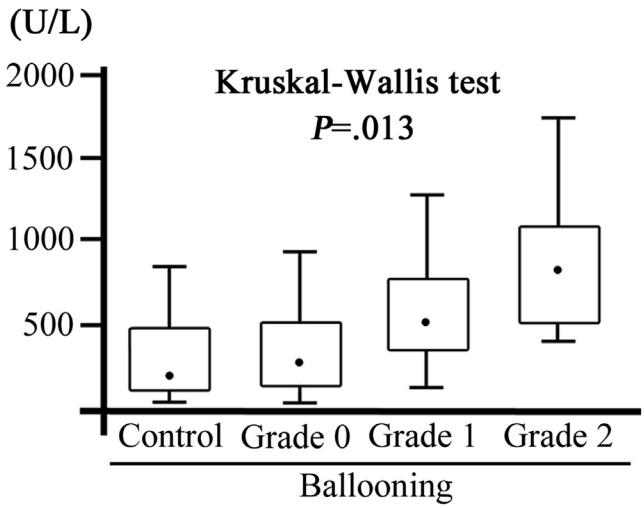
Assessment of Nonalcoholic Steatohepatitis or Nonalcoholic Fatty Liver Disease Activity Score ≥5 in Patients With Nonalcoholic Fatty Liver Disease Using Magnetic Resonance Imaging and Transient Elastography Methods With Serum Keratin 18 and Alanine Transaminase Levels.

We assessed the diagnostic ability of the LSM obtained using MRE combined with MRI-based PDFF (MRE+PDFF) and the LSM obtained using TE including CAP (TE+CAP) regarding NASH and NAS ≥5. For the detection of NASH or NAS ≥5, the AUROC curve for MRI+PDFF and TE+CAP was 0.81 (95% CI: 0.73−0.89) and 0.80 (95% CI: 0.73−0.88), or 0.77 (95% CI: 0.67−0.87), and 0.65 (95% CI: 0.54−0.77), respectively, using ROC analysis (Supplementary Table 3).

Mean serum K18 levels exhibited stepwise increases with increasing histologic severity of ballooned hepatocytes (Supplementary Figure 1; 245 U/L (range, 14−1254 U/L) for control, 388 U/L (range, 21−1965 U/L) for grade 0, 525 U/L (range, 79−3259 U/L) for grade 1, and 867U/L (range 414−1721 U/L) for grade 2; $P = .013$ using the Kruskal-Wallis test). The AUROC curve regarding the performance of the serum K18 levels in distinguishing the grade of

ballooned hepatocyte is shown in Supplementary Table 4. Additionally, the cutoff level and the sensitivity, specificity, positive predictive value, and negative predictive value for each ballooned hepatocyte grade are also detailed in Supplementary Table 4.

Finally, we analyzed the diagnostic ability of serum K18 and alanine transaminase (ALT) levels with MRE+PDFF and TE+CAP in relation to NASH and NAS $\geq$5. The AUROC curve results indicate that MRE+PDFF+K18 or ALT was significantly superior to TE+CAP+K18 or ALT in the diagnosis of NAS $\geq$5 (Supplementary Tables 5 and 6). However, we found that the addition of serum K18 or ALT levels did not improve the diagnostic ability of MRE+PDFF and TE+CAP significantly concerning NASH and NAS $\geq$5 (Supplementary Tables 5 and 6).



Supplementary Figure 1. Relationship between serum K18 levels and ballooned hepatocyte grade in patients with NAFLD. (*A*) A steady stepwise increase in serum K18 levels was observed with increasing severity of ballooned hepatocytes (Kruskal-Wallis test; $P = .013$). The vertical axis represents serum K18 levels in U/L and the horizontal axis represents each ballooned hepatocyte grade. The graph represents the interquartile range (*box*), median (*dot*), and range (*lines*) of the serum K18 levels.

Supplementary Table 1. The Cases in Which it Was Not Possible to Measure LSM and CAP Using TE

Case (n = 15)	BMI (kg/m^2)	Ascites	Interposition of intestine or colon
1	29.4	-	
2	31.1	-	-
3	28.5	-	-
4	29.4	-	-
5	32.2	-	-
6	33.2	-	-
7	27.8	-	-
8	30.2	-	-
9	28.9	-	-
10	29.3	-	-
11	27.0	+	-
12	26.5	+	-
13	27.2	+	-
14	25.9	-	+
15	26.2	-	+

Supplementary Table 2. Multiple Regression Analysis of Histologic Parameters Associated With Increased LSM for MRE or TE and Increased Fat Accumulation for PDFF for MRI or CAP for TE

Case (n = 15)	BMI (kg/m^2)	Ascites	Interposition of intestine or colon
1	29.4	-	
2	31.1	-	-
3	28.5	-	-
4	29.4	-	-
5	32.2	-	-
6	33.2	-	-
7	27.8	-	-
8	30.2	-	-
9	28.9	-	-
10	29.3	-	-
11	27.0	+	-
12	26.5	+	-
13	27.2	+	-
14	25.9	-	+
15	26.2	-	+

Supplementary Table 3. Diagnostic Accuracy of MRE Combined with MRI-Based PDFF and TE Combined TE-Based CAP in Detecting NASH and NAS >5

	MRE+PDFF (n = 127)		TE+CAP (n = 127)		
	AUROC	95%CI	AUROC	95%CI	vs MRE+PDFF *P* value
NASH	0.81	0.73–0.89	0.80	0.73–0.88	0.475
NAS ≥ 5	0.77	0.67–0.87	0.65	0.54–0.77	0.045[a]

[a]MRE+PDFF was significantly superior to TE+CAP.

Supplementary Table 4. Diagnostic Accuracy of Serum K18 Levels in Detecting Each Grade of Ballooned Hepatocyte

	K18 (n = 127)						
Ballooned hepatocyte grade	Cut off level (U/L)	AUROC	95%CI	Se	Sp	PPV	NPV
≥ 1	354	0.67	0.53–0.81	64.4	68.9	86.2	39.2
≥ 2	414	0.80	0.67–0.92	100.0	54.7	17.2	100.0

Supplementary Table 5. Diagnostic Accuracy of MRE+PDFF or TE+CAP Combined With Serum K18 Levels Determined by Detecting NASH and NAS >5

	MRE+PDFF+K18 (n = 127)				TE+CAP+K18 (n = 127)				
Parameters	AUROC	95% CI	*P* value	*P* value (MRE+PDFF vs MRE+PDFF+K18)	AUROC	95% CI	*P* value	*P* value (TE+CAP vs TE+CAP+K18)	*P* value (MRE+PDFF+K18 vs TE+CAP+K18)
NASH	0.82	0.73–0.92	<.001	0.543	0.82	0.75–0.89	<.001	0.539	0.668
NAS ≥ 5	0.78	0.68–0.88	.002	0.717	0.66	0.54–0.78	.048	0.684	0.049[a]

[a]MRE+PDFF+K18 was significantly superior to TE+CAP+K18.

Supplementary Table 6. Diagnostic Accuracy of MRE+PDFF or TE+CAP Combined With Serum ALT Levels Determined by Detecting NASH and NAS >5

	MRE+PDFF+ALT (n = 127)				TE+CAP+ALT (n = 127)				
Parameters	AUROC	95% CI	*P* value	*P* value (MRE+PDFF vs MRE+PDFF+ALT)	AUROC	95% CI	*P* value	*P* value (TE+CAP vs TE+CAP+ALT)	*P* value (MRE+PDFF+ALT vs TE+CAP+ALT)
NASH	0.82	0.74–0.91	<.001	0.485	0.80	0.72–0.88	<.001	0.515	0.325
NAS ≥ 5	0.77	0.67–0.87	.004	0.722	0.67	0.56–0.78	.043	0.483	0.047[a]

[a]MRE+PDFF+ALT was significantly superior to TE+CAP+ALT.

Supplementary Table 7. The Cost of MRI and TE Method in Several Countries

	Cost *(US dollars)*	
Countries	MRE	TE
Japan	130	65
USA	2871	137
United Kingdom	335	216
France	363	29